```
XX
PI   Goli SK,  Hillman JL;
XX
DR   WPI; 1998-447238/38.
DR   P-PSDB; AAW61538.
XX
PT   New human late embryogenesis abundant-like protein - useful to treat
PT   disordered cell proliferation and growth, or to induce tissue
PT   regeneration
XX
PS   Claim 1; Fig 1; 57pp; English.
XX
CC   This sequence encodes the human LEA (late embryogenesis abundant)-motif
CC   protein (HuLEAP) of the invention. Cells containing the DNA are used to
CC   produce recombinant HuLEAP. HuLEAP and its agonists (optionally expressed
CC   from gene therapy vectors) are used to treat disorders that involve
CC   abnormal cellular differentiation or growth, e.g. hypoaldosteronism,
CC   Addison's disease, hypothyroidism, colorectal polyps, gastric and
CC   duodenal ulcers, haematopoietic, lymphoid or many other forms of cancer.
CC   Antagonists of HuLEAP, e.g. Ab, antisense or untranslatable sense
CC   sequences or ribozymes, are used to induce cell proliferation,
CC   particularly in tissues that do not regenerate under normal conditions,
CC   e.g. in cases or trauma or degenerative diseases, and for in vivo/in
CC   vitro stimulation of skin growth for grafting. Ab can be used as an
CC   antagonist, as carrier to deliver agents to HuLEAP-expressing cells, as
CC   immunoassay for (diagnostic) detection of HuLEAP, in competitive drug
CC   screens and to purify HuLEAP from natural sources. Fragments of the DNA
CC   can be used diagnostically (in hybridisation or amplification tests), for
CC   mapping genomic sequences and in drug screening.
XX
SQ   Sequence 858 BP; 191 A; 258 C; 245 G; 164 T; 0 other;

  Query Match            100.0%;  Score 858;  DB 19;  Length 858;
  Best Local Similarity  100.0%;  Pred. No. 4.9e-206;
  Matches  858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60

Qy         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120

Qy        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180

Qy        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240

Qy        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300

Qy        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360

Qy        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420

Qy        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480

Qy        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540

Qy        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600

Qy        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660

Qy        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720

Qy        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780

Qy        781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840

Qy        841 CCCTCCCTTCATTGTACT 858
              ||||||||||||||||||
Db        841 CCCTCCCTTCATTGTACT 858


RESULT 2
AAA46361
ID   AAA46361 standard; DNA; 1549 BP.
XX
AC   AAA46361;
XX
DT   04-SEP-2000  (first entry)
XX
DE   Nucleotide sequence of the gene insert of CINN 1.
XX
KW   Malignant cell; OP2 C2-6; OP7 C3-1; OP9 A4-2; OP11 C1-3; OP11 G2-10;
KW   FAS OP13; FAS OP17; CINN 1; CINN 2; anti-neoplastic drug; malignancy;
KW   cancer; brain cancer; intracranial neoplasm; ss.
XX
OS   Homo sapiens.
XX
PN   WO200028090-A2.
XX
PD   18-MAY-2000.
XX
PF   12-NOV-1999;   99WO-US27005.
XX
PR   12-NOV-1998;   98US-0108120.
PR   27-JUL-1999;   99US-0145640.
XX
PA   (NYXI-) NYXIS INC.
XX
PI   Kroes RA,  Moskal JR,  Jastrow A,  Yamamoto H;
XX
DR   WPI; 2000-376595/32.
XX
PT   Method for ascertaining the propensity of a cell for malignant
PT   phenotype, used for diagnosis of cancers, including those of the lung,
PT   breast and brain  -
XX
PS   Claim 1; Fig 1; 66pp; English.
XX
CC   The specification describes a method for ascertaining the propensity of
CC   a cell for malignant phenotype. The method comprises assaying a cell or
CC   biological sample for a signal indicating the transcription of a nucleic
CC   acid transcript. The transcript is from at least one gene selected from
CC   the genes encoded for by or containing the characteristic nucleic acid
CC   sequences of OP2 C2-6, OP7 C3-1, OP9 A4-2, OP11 C1-3, OP11 G2-10,
CC   FAS OP13, FAS OP17, CINN 1 or CINN 2 (AAA46354-62, respectively). Cells
CC   identified using the method can be used for ascertaining the suitability
CC   of an anti-neoplastic drug candidate for efficacy in treating a
CC   malignancy. The method can be used in the diagnosis of many types of
CC   cancers, including cancers of the breast, prostate, colon and lung. It
CC   may also be used for diagnosis and treatment of brain cancers, including
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 9, 2003, 10:32:57 ; Search time 326 Seconds
(without alignments)
7104.656 Million cell updates/sec

Title: US-09-892-316-2
Perfect score: 858
Sequence: 1 GCGGTGGTGACTGAGCTACG.....TGCCCTCCCTTCATTGTACT 858

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	858	100.0	858	19	AAV45175	Human LEA-motif de
2	853.6	99.5	1549	21	AAA46361	Nucleotide sequenc
c 3	524.4	61.1	1149	22	ABA09672	Human bone marrow
4	436.2	50.8	601	21	AAA16360	Human colon cancer
c 5	382.6	44.6	1300	23	AAS66254	DNA encoding novel
c 6	380.2	44.3	620	22	ABA09588	Human bone marrow
c 7	374.2	43.6	1298	23	AAS65873	DNA encoding novel
8	357	41.6	501	23	AAS73953	DNA encoding novel
c 9	314.6	36.7	384	20	AAV86069	EST clone D14. Ho
10	289.4	33.7	406	25	ABX47759	Bovine EST associa
c 11	251	29.3	281	24	ABN94115	Gene #613 used to
12	188.6	22.0	983	24	ABN73957	Bovine embryonic g
c 13	184.2	21.5	416	22	AAH83599	Human ovarian tumo
c 14	182	21.2	316	22	AAS24954	Human ovarian PCR-
15	162.4	18.9	930	24	ABN73958	Bovine embryonic g
16	148.2	17.3	257	25	ABX54860	Bovine EST associa
17	132.8	15.5	948	24	ABN74707	Bovine embryonic g
18	86.6	10.1	475	23	AAS73954	DNA encoding novel
19	69.2	8.1	1099	23	ABL04817	Drosophila melanog
c 20	69.2	8.1	3547	23	ABL04816	Drosophila melanog
21	49.4	5.8	2765	21	AAA96707	Reporter gene YLR1
22	48.8	5.7	709	20	AAZ15183	Human gene express
23	40.4	4.7	511	24	ABN73443	Bovine embryonic g
24	40.2	4.7	4804	21	AAZ46086	cDNA encoding a re
25	40	4.7	4457	23	ABL06791	Drosophila melanog
c 26	40	4.7	9051	23	ABL06790	Drosophila melanog
c 27	39.4	4.6	404	25	ABZ55322	Aspergillus oryzae
28	39.4	4.6	1588	22	AAH26304	Spider silk protei
29	39.2	4.6	1665	23	ABL28565	Drosophila melanog
30	39.2	4.6	3633	18	AAT97611	Mouse E2A-binding
c 31	39.2	4.6	3665	23	ABL28564	Drosophila melanog
32	39.2	4.6	3728	15	AAQ58708	Mouse OSP-5 coding
33	39	4.5	1626	22	AAI59140	Human polynucleoti
34	39	4.5	1685	24	ABL50340	Human cancer cell
35	39	4.5	2556	22	AAI60926	Human polynucleoti
36	39	4.5	3048	24	ABZ35239	Human gene express
37	38.8	4.5	2166	24	ABL55863	Human oxidoreducta
38	38.8	4.5	50000	21	AAA64139	Nucleotide sequenc
39	38.6	4.5	1566	23	ABL13201	Drosophila melanog
c 40	38.6	4.5	2927	23	ABL13202	Drosophila melanog
c 41	38.6	4.5	3636	23	ABL13200	Drosophila melanog
42	38.4	4.5	786	20	AAX39713	Gastric cancer ass
43	38.4	4.5	927	20	AAX39716	Gastric cancer ass
44	38.4	4.5	1011	24	ABZ34915	Human gene express
45	38.4	4.5	1187	24	ABZ35437	Human gene express

ALIGNMENTS

RESULT 1
AAV45175
ID AAV45175 standard; DNA; 858 BP.
XX
AC AAV45175;
XX
DT 29-OCT-1998 (first entry)
XX
DE Human LEA-motif developmental protein coding sequence.
XX
KW LEA-motif developmental protein; human; HuLEAP; tissue regeneration;
KW abnormal cellular differentiation; hypoaldosteronism; Addison's disease;
KW hypothyroidism; colorectal polyps; duodenal ulcer; cancer; therapy;
KW late embryogenesis abundant protein; cell proliferation induction; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 130..790
FT /*tag= a
FT /product= HuLEAP
XX
PN WO9835041-A1.
XX
PD 13-AUG-1998.
XX
PF 05-FEB-1998; 98WO-US02470.
XX
PR 06-FEB-1997; 97US-0796676.
XX
PA (INCY-) INCYTE PHARM INC.

PT New human late embryogenesis abundant-like protein - useful to treat
PT disordered cell proliferation and growth, or to induce tissue
PT regeneration
XX
PS Claim 5; Fig 1; 57pp; English.
XX
CC This sequence is the human LEA (late embryogenesis abundant)-motif
CC protein (HuLEAP) of the invention. Cells containing the DNA are used to
CC produce recombinant HuLEAP. HuLEAP and its agonists (optionally expressed
CC from gene therapy vectors) are used to treat disorders that involve
CC abnormal cellular differentiation or growth, e.g. hypoaldosteronism,
CC Addison's disease, hypothyroidism, colorectal polyps, gastric and
CC duodenal ulcers, haematopoietic, lymphoid or many other forms of cancer.
CC Antagonists of HuLEAP, e.g. Ab, antisense or untranslatable sense
CC sequences or ribozymes, are used to induce cell proliferation,
CC particularly in tissues that do not regenerate under normal conditions,
CC e.g. in cases or trauma or degenerative diseases, and for in vivo/in
CC vitro stimulation of skin growth for grafting. Ab can be used as an
CC antagonist, as carrier to deliver agents to HuLEAP-expressing cells, as
CC immunoassay for (diagnostic) detection of HuLEAP, in competitive drug
CC screens and to purify HuLEAP from natural sources. Fragments of the DNA
CC can be used diagnostically (in hybridisation or amplification tests), for
CC mapping genomic sequences and in drug screening.
XX
SQ Sequence 219 AA;

Query Match 100.0%; Score 1130; DB 19; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.5e-106;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60

Qy         61 RMPRWAERLFPANVAHSVYVLEDSIVDPQNQTMTTFTWNINHARLMVVEERCVYCVNSDN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RMPRWAERLFPANVAHSVYVLEDSIVDPQNQTMTTFTWNINHARLMVVEERCVYCVNSDN 120

Qy        121 SGWTEIRREAWVSSSLFGVSRAVQEFGLARFKSNVTKTMKGFEYILAKLQGEAPSKTLVE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGWTEIRREAWVSSSLFGVSRAVQEFGLARFKSNVTKTMKGFEYILAKLQGEAPSKTLVE 180

Qy        181 TAKEAKEKAKETALAATEKAKDLASKAATKKQQQQQQFV 219
              |||||||||||||||||||||||||||||||||||||||
Db        181 TAKEAKEKAKETALAATEKAKDLASKAATKKQQQQQQFV 219
```

RESULT 2
ABG09766
ID ABG09766 standard; Protein; 167 AA.
XX
AC ABG09766;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9757.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS73953.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 40125; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 167 AA;

Query Match 64.3%; Score 726.5; DB 22; Length 167;
Best Local Similarity 69.6%; Pred. No. 1.1e-65;
Matches 149; Conservative 8; Mismatches 10; Indels 47; Gaps 2;

```
Qy          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60
              ||||:||| |||||||||  |||||| ||  :: |||||||||
Db          1 MVKYYLGQGVLRSSWDQVVPAFWQRYRNPVQRY-LTEDIVHRE---------------- 42

Qy         61 RMPRWAERLFPANVAHSVYVLEDSIVDPQNQTMTTFTWNINHARLMVVEERCVYCVNSDN 120
                                              ||||||||||||||| ||||||||||||||
Db         43 ----------------------------NQTMTTFTWNINHARPMVVEERCVYCVNSDN 73

Qy        121 SGWTEIRREAWVSSSLFGVSRAVQEFGLARFKSNVTKTMKGFEYILAKLQGEAPSKTLVE 180
              |||||||||||||||| |||||| |||||:|||||||:|||||||||||||| ||||::
Db         74 SGWTEIRREAWVSSSLCGVSRAVWEFGLAQFKSNVTKTLKGFEYILAKLQGEAASKTLIK 133

Qy        181 TAKEAKEKAKETALAATEKAKDLASKAATKKQQQ 214
              ||:|||||||||||||||||||||||||||||
Db        134 TAREAKEKAKETALAATEKAKDLASKAATKKQQQ 167
```

RESULT 3
ABB12426
ID ABB12426 standard; Protein; 99 AA.
XX
AC ABB12426;
XX
DT 15-JAN-2002 (first entry)
XX
DE Human bone marrow expressed protein SEQ ID NO: 265.
XX
KW Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnerary;
KW antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
KW antiulcer; fungicide; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; cerebroprotective; antianaemic; infection;
KW nervous system disorder; autoimmune disorder; inflammation; allergy.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 12:58:13 ; Search time 73 Seconds
(without alignments)
476.180 Million cell updates/sec

Title: US-09-892-316-1
Perfect score: 1130
Sequence: 1 MVKYFLGQSVLRSSWDQVFA.....AKDLASKAATKKQQQQQQFV 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1130	100.0	219	19	AAW61538	Human LEA-motif de
2	726.5	64.3	167	22	ABG09766	Novel human diagno
3	339.5	30.0	99	22	ABB12426	Human bone marrow
4	245.5	21.7	98	22	ABG02067	Novel human diagno
5	229	20.3	171	22	ABB59225	Drosophila melanog
6	217.5	19.2	82	22	ABG01686	Novel human diagno
7	200.5	17.7	194	21	AAB03814	Human vesicle traf
8	200.5	17.7	194	21	AAY49959	Human vesicle traf
9	200.5	17.7	194	22	AAB93664	Human protein sequ
10	198	17.5	215	22	ABB64259	Drosophila melanog
11	179.5	15.9	183	21	AAG17835	Arabidopsis thalia
12	179.5	15.9	222	21	AAG17834	Arabidopsis thalia
13	148	13.1	75	22	ABG09767	Novel human diagno
14	143	12.7	117	23	ABP03962	Human ORFX protein
15	142.5	12.6	164	23	ABP69539	Human polypeptide
16	139.5	12.3	715	24	ABU03523	Angiogenesis-assoc
17	137	12.1	696	24	AAE34445	Human lipid-associ
18	125	11.1	659	22	ABB64495	Drosophila melanog
19	114	10.1	177	22	AAB36623	Human FLEXHT-45 pr
20	111	9.8	138	21	AAG20424	Arabidopsis thalia
21	108	9.6	387	22	ABG27715	Novel human diagno
22	98	8.7	59	21	AAG02195	Human secreted pro
23	89	7.9	674	21	AAY91922	Candida albicans S
24	88.5	7.8	2101	21	AAY49936	Human NuMA protein
25	87	7.7	140	22	ABG27711	Novel human diagno
26	85.5	7.6	417	22	ABG17962	Novel human diagno
27	85.5	7.6	433	18	AAW30256	Zuotin. Saccharom
28	85.5	7.6	433	20	AAY32954	Human zuotin prote
29	85.5	7.6	433	22	AAG70739	S cerevisiae apopt
30	83.5	7.4	2101	15	AAR47173	Sequence of the in
31	83.5	7.4	2101	22	AAG65799	Nuclear mitotic ap
32	83.5	7.4	2207	22	AAU32041	Novel human secret
33	83.5	7.4	2356	24	ABJ25846	Aspergillus fumiga
34	83.5	7.4	2816	22	AAU68572	Human novel cytoki
35	83.5	7.4	3170	24	ABJ26446	Aspergillus fumiga
36	83	7.3	7201	22	ABB71136	Drosophila melanog
37	82.5	7.3	260	23	ABB48773	Listeria monocytog
38	82.5	7.3	427	18	AAW34211	Polyketide synthas
39	82.5	7.3	904	22	ABB71477	Drosophila melanog
40	82.5	7.3	904	23	ABG70014	Larval viability a
41	81.5	7.2	780	24	AAE34441	Human lipid-associ
42	81	7.2	349	21	AAY98038	Yeast SYR2, confer
43	81	7.2	426	19	AAW55812	Streptomyces roseo
44	80.5	7.1	1354	22	ABB60329	Drosophila melanog
45	80.5	7.1	1637	23	ABB79896	Human kinesin moto

ALIGNMENTS

RESULT 1
AAW61538
ID AAW61538 standard; Protein; 219 AA.
XX
AC AAW61538;
XX
DT 29-OCT-1998 (first entry)
XX
DE Human LEA-motif developmental protein.
XX
KW LEA-motif developmental protein; human; HuLEAP; tissue regeneration;
KW abnormal cellular differentiation; hypoaldosteronism; Addison's disease;
KW hypothyroidism; colorectal polyps; duodenal ulcer; cancer; therapy;
KW late embryogenesis abundant protein; cell proliferation induction.
XX
OS Homo sapiens.
XX
PN WO9835041-A1.
XX
PD 13-AUG-1998.
XX
PF 05-FEB-1998; 98WO-US02470.
XX
PR 06-FEB-1997; 97US-0796676.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Goli SK, Hillman JL;
XX
DR WPI; 1998-447238/38.
DR N-PSDB; AAV45175.
XX

```
RC   TISSUE=Brain, Eye, Ovary, and Placenta;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be important for the development of vital and
CC       immunocompetent organs.
CC   -!- TISSUE SPECIFICITY: Highly expressed in fetal liver; less
CC       expressed in fetal brain, lung, and kidney. At the adult stage,
CC       expression is drastically reduced in the liver but highly
CC       expressed in the spleen, brain, lung, lymph nodes and peripheral
CC       blood leukocytes.
CC   -!- SIMILARITY: Contains 1 PRELI/MSF1 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF201925; AAF09255.1; -.
DR   EMBL; AF151864; AAD34101.1; -.
DR   EMBL; AF153607; AAD41089.1; -.
DR   EMBL; AF111112; AAF27195.1; -.
DR   EMBL; AF112203; AAF17191.1; -.
DR   EMBL; BC000007; AAH00007.1; -.
DR   EMBL; BC007268; AAH07268.1; -.
DR   EMBL; BC008307; AAH08307.1; -.
DR   EMBL; BC008866; AAH08866.1; -.
DR   EMBL; BC013733; AAH13733.1; -.
DR   EMBL; BC013748; AAH13748.1; -.
DR   MIM; 605733; -.
DR   GO; GO:0007275; P:development; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   InterPro; IPR006797; MSF1.
DR   Pfam; PF04707; MSF1; 1.
DR   PROSITE; PS50904; PRELI_MSF1; 1.
FT   DOMAIN        1    174       PRELI/MSF1.
FT   CONFLICT    127    127       R -> H (IN REF. 1).
FT   CONFLICT    142    142       A -> T (IN REF. 5).
FT   CONFLICT    171    181       MISSING (IN REF. 5).
SQ   SEQUENCE   219 AA;  25181 MW;  CCDBF54FA573509F CRC64;

  Query Match             100.0%;  Score 1130;  DB 1;  Length 219;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-86;
  Matches  219;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60

Qy         61 RMPRWAERLFPANVAHSVYVLEDSIVDPQNQTMTTFTWNINHARLMVVEERCVYCVNSDN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RMPRWAERLFPANVAHSVYVLEDSIVDPQNQTMTTFTWNINHARLMVVEERCVYCVNSDN 120

Qy        121 SGWTEIRREAWVSSSLFGVSRAVQEFGLARFKSNVTKTMKGFEYILAKLQGEAPSKTLVE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGWTEIRREAWVSSSLFGVSRAVQEFGLARFKSNVTKTMKGFEYILAKLQGEAPSKTLVE 180

Qy        181 TAKEAKEKAKETALAATEKAKDLASKAATKKQQQQQQFV 219
              |||||||||||||||||||||||||||||||||||||||
Db        181 TAKEAKEKAKETALAATEKAKDLASKAATKKQQQQQQFV 219


RESULT 2
PX19_MOUSE
ID   PX19_MOUSE     STANDARD;      PRT;   217 AA.
AC   Q8R107;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Px19-like protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 1 PRELI/MSF1 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC025859; AAH25859.1; -.
DR   InterPro; IPR006797; MSF1.
DR   Pfam; PF04707; MSF1; 1.
DR   PROSITE; PS50904; PRELI_MSF1; 1.
FT   DOMAIN        1    174       PRELI/MSF1.
SQ   SEQUENCE   217 AA;  24959 MW;  8D909AB939B6D851 CRC64;

  Query Match             97.2%;  Score 1098;  DB 1;  Length 217;
  Best Local Similarity   98.1%;  Pred. No. 3.9e-83;
  Matches  211;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVKYFLGQSVLRSSWDQVFAAFWQRYPNPYSKHVLTEDIVHREVTPDQKLLSRRLLTKTN 60

Qy         61 RMPRWAERLFPANVAHSVYVLEDSIVDPQNQTMTTFTWNINHARLMVVEERCVYCVNSDN 120
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 9, 2003, 12:59:53 ; Search time 23 Seconds
(without alignments)
447.776 Million cell updates/sec

Title: US-09-892-316-1
Perfect score: 1130
Sequence: 1 MVKYFLGQSVLRSSWDQVFA.....AKDLASKAATKKQQQQQQFV 219

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1130	100.0	219	1	PX19_HUMAN	Q9y255 homo sapien
2	1098	97.2	217	1	PX19_MOUSE	Q8r107 mus musculu
3	958	84.8	215	1	PX19_CHICK	Q90673 gallus gall
4	207.5	18.4	230	1	MSF1_YEAST	P35200 saccharomyc
5	201.5	17.8	195	1	CT45_MOUSE	Q9cyy7 mus musculu
6	200.5	17.7	194	1	CT45_HUMAN	Q9y3b1 homo sapien
7	139.5	12.3	715	1	S141_HUMAN	Q92503 homo sapien
8	126	11.2	743	1	YN02_CAEEL	Q03606 caenorhabdi
9	107	9.5	747	1	VIVC_BPT7	P03725 bacteriopha
10	92	8.1	422	1	GAS7_RAT	O55148 rattus norv
11	91.5	8.1	421	1	GAS7_MOUSE	Q60780 mus musculu
12	90	8.0	644	1	YEN9_YEAST	P40032 saccharomyc
13	85.5	7.6	412	1	GAS7_HUMAN	O60861 homo sapien
14	85.5	7.6	433	1	ZUO1_YEAST	P32527 saccharomyc
15	85	7.5	354	1	VF11_VACCC	P21052 vaccinia vi
16	84	7.4	822	1	FER_HUMAN	P16591 homo sapien
17	84	7.4	878	1	MSH4_YEAST	P40965 saccharomyc
18	83	7.3	359	1	SERC_BACSU	P80862 bacillus su
19	83	7.3	1787	1	CHD3_CAEEL	Q22516 caenorhabdi
20	82	7.3	2832	1	NDVB_RHIME	P20471 rhizobium m
21	81	7.2	349	1	SUR2_YEAST	P38992 saccharomyc
22	80.5	7.1	687	1	TRA_HAEIN	O05069 haemophilus
23	80.5	7.1	700	1	ADB1_YEAST	P27351 saccharomyc
24	80.5	7.1	1959	1	MYH9_CHICK	P14105 gallus gall
25	80	7.1	372	1	TOLA_HAEIN	P44678 haemophilus
26	80	7.1	1403	1	VG22_HSVI1	Q00105 ictalurid h
27	79.5	7.0	121	1	RS24_PYRAE	Q8zv65 pyrobaculum
28	79	7.0	337	1	VF11_VACCP	P29888 vaccinia vi
29	79	7.0	426	1	GUNZ_ERWCH	P07103 erwinia chr
30	78.5	6.9	280	1	LE76_BRANA	P13934 brassica na
31	78.5	6.9	967	1	CAP2_MAIZE	P51059 zea mays (m
32	78	6.9	224	1	Y364_MYCGE	P47604 mycoplasma
33	78	6.9	471	1	K1CN_HUMAN	P02533 homo sapien
34	78	6.9	540	1	C314_DROME	Q9vuf8 drosophila
35	77.5	6.9	203	1	VATE_DESSY	O06501 desulfuroco
36	77.5	6.9	250	1	H11_DROVI	Q24704 drosophila
37	77.5	6.9	250	1	H12_DROVI	Q94555 drosophila
38	77.5	6.9	269	1	NBPX_MOUSE	Q9dbl7 mus musculu
39	77.5	6.9	376	1	G4P1_YEAST	P46672 saccharomyc
40	77.5	6.9	483	1	M6_STRPY	P08089 streptococc
41	77.5	6.9	565	1	MOBA_BIFLO	Q8gn32 bifidobacte
42	77.5	6.9	570	1	ETFD_ACICA	P94132 acinetobact
43	77	6.8	124	1	RS16_RHIME	Q92143 rhizobium m
44	77	6.8	539	1	GSHB_ARATH	P46416 arabidopsis
45	77	6.8	577	1	SYR_VIBPA	Q87rd6 vibrio para

ALIGNMENTS

RESULT 1
PX19_HUMAN
ID PX19_HUMAN STANDARD; PRT; 219 AA.
AC Q9Y255; Q9UI13; Q9UJS9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Px19-like protein (25 kDa protein of relevant evolutionary and
DE lymphoid interest) (PRELI) (CGI-106) (SBBI12).
GN PRELI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=20247012; PubMed=10784606;
RA Guzman-Rojas L., Sims J.C., Rangel R., Guret C., Sun Y., Alcocer J.M.,
RA Martinez-Valdez H.;
RT "PRELI, the human homologue of the avian px19, is expressed by
RT germinal center B lymphocytes.";
RL Int. Immunol. 12:607-612(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Dermal papilla;
RA Kim M.K., Kim Y.H., Seo J.M., Lee H.M., Chung H.J., Sohn M.Y.,
RA Hwang S.Y., Im S.U., Jung E.J., Lee J.H., Kim J.C.;
RT "A catalogue of genes in the human dermal papilla cells as identified
RT by expressed sequence tags.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Zhang W., He L., Wan T., Zhu X., Cao X.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal gland;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN [6]
RP SEQUENCE FROM N.A.

```
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120

Qy         121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180

Qy         181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240

Qy         241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300

Qy         301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360

Qy         361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420

Qy         421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480

Qy         481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540

Qy         541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600

Qy         601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660

Qy         661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720

Qy         721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780

Qy         781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840

Qy         841 CCCTCCCTTCATTGTACT 858
                ||||||||||||||||||
Db         841 CCCTCCCTTCATTGTACT 858


RESULT 2
US-09-213-391-2
; Sequence 2, Application US/09213391
; Patent No. 6281190
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  HUMAN LEA-MOTIF DEVELOPMENTAL
;    TITLE OF INVENTION:  PROTEIN
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/213,391
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/796,676
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0213 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  858 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-213-391-2

  Query Match             100.0%;  Score 858;  DB 3;  Length 858;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-214;
  Matches  858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60

Qy          61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120

Qy         121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180

Qy         181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240

Qy         241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300

Qy         301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360

Qy         361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420

Qy         421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480

Qy         481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540

Qy         541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600

Qy         601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 9, 2003, 11:04:23 ; Search time 78 Seconds
(without alignments)
4855.214 Million cell updates/sec

Title: US-09-892-316-2
Perfect score: 858
Sequence: 1 GCGGTGGTGACTGAGCTACG.....TGCCCTCCCTTCATTGTACT 858

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	858	100.0	858	2	US-08-796-676-2	Sequence 2, Appli
	2	858	100.0	858	3	US-09-213-391-2	Sequence 2, Appli
	3	853.6	99.5	1549	3	US-09-438-938-8	Sequence 8, Appli
	4	853.6	99.5	1549	4	US-09-668-885A-8	Sequence 8, Appli
	5	436.2	50.8	601	3	US-09-385-982-365	Sequence 365, App
c	6	39.2	4.6	642	4	US-09-252-991A-15522	Sequence 15522, A
	7	39.2	4.6	822	4	US-09-252-991A-15264	Sequence 15264, A
c	8	39.2	4.6	993	4	US-09-252-991A-15411	Sequence 15411, A
c	9	39.2	4.6	1644	4	US-09-252-991A-15473	Sequence 15473, A
	10	39.2	4.6	3728	1	US-08-111-939-1	Sequence 1, Appli
	11	39	4.5	1626	4	US-09-620-312D-1033	Sequence 1033, Ap
c	12	38.6	4.5	7218	1	US-08-232-463-14	Sequence 14, Appl
	13	38	4.4	1374	4	US-09-252-991A-8560	Sequence 8560, Ap
	14	38	4.4	2463	4	US-09-252-991A-8582	Sequence 8582, Ap
c	15	38	4.4	2748	4	US-09-252-991A-8662	Sequence 8662, Ap
	16	37.6	4.4	1392	4	US-09-252-991A-7744	Sequence 7744, Ap
c	17	37.6	4.4	1905	4	US-09-252-991A-7829	Sequence 7829, Ap
	18	37.6	4.4	43280	2	US-08-804-227C-1	Sequence 1, Appli
	19	37.2	4.3	3202	4	US-09-287-354-1	Sequence 1, Appli
c	20	37.2	4.3	6529	3	US-08-789-329C-1	Sequence 1, Appli
c	21	36.4	4.2	1514	4	US-09-162-524-4	Sequence 4, Appli
	22	36.4	4.2	2988	4	US-09-996-243-325	Sequence 325, App
c	23	36.2	4.2	2571	4	US-09-984-880-1	Sequence 1, Appli
c	24	36.2	4.2	20966	4	US-09-984-880-3	Sequence 3, Appli
c	25	36.2	4.2	28804	2	US-08-592-874-1	Sequence 1, Appli
c	26	36.2	4.2	28804	3	US-09-096-942-2	Sequence 2, Appli
c	27	36.2	4.2	28804	3	US-09-096-867-2	Sequence 2, Appli
	28	36	4.2	3935	4	US-09-060-482-1	Sequence 1, Appli
	29	35.8	4.2	289	3	US-09-007-005-17	Sequence 17, Appl
	30	35.8	4.2	289	3	US-09-244-796-17	Sequence 17, Appl
	31	35.8	4.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	32	35.8	4.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	33	35.4	4.1	550	3	US-08-705-771-1	Sequence 1, Appli
c	34	35.4	4.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	35	35.4	4.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	36	35	4.1	402	4	US-09-252-991A-7998	Sequence 7998, Ap
c	37	35	4.1	1236	4	US-09-252-991A-8061	Sequence 8061, Ap
	38	35	4.1	1671	4	US-09-252-991A-8261	Sequence 8261, Ap
	39	35	4.1	1767	4	US-09-016-434-1488	Sequence 1488, Ap
	40	35	4.1	2556	4	US-09-252-991A-8202	Sequence 8202, Ap
	41	35	4.1	3205	3	US-09-046-894-34	Sequence 34, Appl
	42	35	4.1	3205	4	US-09-220-132-165	Sequence 165, App
	43	34.8	4.1	516	4	US-09-252-991A-15740	Sequence 15740, A
	44	34.8	4.1	705	3	US-08-998-416-1059	Sequence 1059, Ap
c	45	34.8	4.1	768	4	US-09-252-991A-15652	Sequence 15652, A

ALIGNMENTS

RESULT 1
US-08-796-676-2
; Sequence 2, Application US/08796676
; Patent No. 5858712
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: HUMAN LEA-MOTIF DEVELOPMENTAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/796,676
; FILING DATE: Filed Herewith
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0213 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-796-676-2

Query Match 100.0%; Score 858; DB 2; Length 858;
Best Local Similarity 100.0%; Pred. No. 4.8e-214;
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCCGGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60

Qy         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120
```

```
          http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0DJ010CB08QP1.
FEATURES             Location/Qualifiers
     source          1. .1062
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ010YD15"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      237 a    302 c    307 g    211 t      5 others
ORIGIN

  Query Match           100.0%;  Score 858;  DB 9;  Length 1062;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-202;
  Matches  858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 155

Qy         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 215

Qy        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 275

Qy        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 335

Qy        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 395

Qy        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 455

Qy        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 515

Qy        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 575

Qy        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 635

Qy        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 695

Qy        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 755

Qy        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        756 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 815

Qy        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        816 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 875

Qy        781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        876 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 935

Qy        841 CCCTCCCTTCATTGTACT 858
              ||||||||||||||||||
Db        936 CCCTCCCTTCATTGTACT 953


RESULT 2
BC018904
LOCUS       BC018904               1252 bp    mRNA    linear   HTC 07-DEC-2001
DEFINITION  Homo sapiens, clone IMAGE:3957371, mRNA.
ACCESSION   BC018904
VERSION     BC018904.1  GI:17403027
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1252)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 15 Row: d Column: 4
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, GenomeScan gene prediction
            This clone has the following problem: no 5' EST match.
FEATURES             Location/Qualifiers
     source          1. .1252
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3957371"
                     /tissue_type="Placenta, choriocarcinoma"
                     /clone_lib="NIH_MGC_21"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
BASE COUNT      285 a    348 c    362 g    257 t
ORIGIN

  Query Match           100.0%;  Score 858;  DB 11;  Length 1252;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-202;
  Matches  858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 9, 2003, 11:03:38 ; Search time 2592 Seconds
(without alignments)
8045.232 Million cell updates/sec

Title: US-09-892-316-2
Perfect score: 858
Sequence: 1 GCGGTGGTGACTGAGCTACG.....TGCCCTCCCTTCATTGTACT 858

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	858	100.0	1062	9	AL559031	AL559031 AL559031
2	858	100.0	1252	11	BC018904	BC018904 Homo sapi
3	857.6	100.0	1201	9	AL559788	AL559788 AL559788
4	832.8	97.1	967	12	BM549485	BM549485 AGENCOURT
5	830.8	96.8	1060	12	BM550706	BM550706 AGENCOURT
6	830	96.7	1018	9	AL583696	AL583696 AL583696
7	827.6	96.5	1201	9	AL529652	AL529652 AL529652
8	826.6	96.3	1030	12	BQ062861	BQ062861 AGENCOURT
9	825	96.2	945	13	BU175909	BU175909 AGENCOURT
10	824.8	96.1	883	13	BU542503	BU542503 AGENCOURT
11	822.2	95.8	938	12	BM541438	BM541438 AGENCOURT
12	815.2	95.0	1054	9	AL558504	AL558504 AL558504
13	813	94.8	911	13	BU166159	BU166159 AGENCOURT
14	813	94.8	953	13	BQ279042	BQ279042 AGENCOURT
15	812.6	94.7	965	13	BQ678206	BQ678206 AGENCOURT
16	809.6	94.4	1172	13	BX404186	BX404186 BX404186
17	806.2	94.0	969	13	BU538405	BU538405 AGENCOURT
18	794.4	92.6	882	13	BU160091	BU160091 AGENCOURT
19	790	92.1	878	13	BQ228147	BQ228147 AGENCOURT
20	788.2	91.9	1065	12	BM541530	BM541530 AGENCOURT
21	780.2	90.9	873	13	BU179062	BU179062 AGENCOURT
22	779.8	90.9	887	13	BQ430303	BQ430303 AGENCOURT
23	773.2	90.1	815	12	BI759007	BI759007 603042487
24	768.2	89.5	880	12	BI258096	BI258096 602970515
25	764.8	89.1	905	13	BQ214286	BQ214286 AGENCOURT
26	763.8	89.0	844	13	BQ218078	BQ218078 AGENCOURT
27	762.2	88.8	914	13	BU552421	BU552421 AGENCOURT
28	756	88.1	889	12	BI759285	BI759285 603042926
29	754.8	88.0	912	13	BU543532	BU543532 AGENCOURT
30	754.6	87.9	892	13	BQ434896	BQ434896 AGENCOURT
31	754	87.9	880	13	BU184058	BU184058 AGENCOURT
32	751.4	87.6	922	13	BQ937888	BQ937888 AGENCOURT
33	750	87.4	766	12	BI259597	BI259597 602968367
34	748.4	87.2	877	9	AL530098	AL530098 AL530098
c 35	743.8	86.7	1201	9	AL529651	AL529651 AL529651
36	743.2	86.6	904	12	BG829541	BG829541 602763843
37	742.6	86.6	909	13	BQ212519	BQ212519 AGENCOURT
38	742.4	86.5	777	12	BI335024	BI335024 602999030
39	739.6	86.2	866	13	BQ425287	BQ425287 AGENCOURT
40	738.4	86.1	843	9	AL583524	AL583524 AL583524
41	737.2	85.9	919	13	BQ225432	BQ225432 AGENCOURT
42	737.2	85.9	1000	14	CD518970	CD518970 AGENCOURT
43	736	85.8	978	12	BQ053570	BQ053570 AGENCOURT
44	735.6	85.7	941	13	BU552441	BU552441 AGENCOURT
45	733.4	85.5	748	12	BI196772	BI196772 602755715

ALIGNMENTS

RESULT 1
AL559031
LOCUS AL559031 1062 bp mRNA linear EST 31-MAY-2003
DEFINITION AL559031 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ010YD15 5-PRIME, mRNA sequence.
ACCESSION AL559031
VERSION AL559031.2 GI:31283164
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12904128.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3031.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ010CB08QP1&cluster=3031.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :

```
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="unknown"
BASE COUNT      191 a    258 c    245 g    164 t
ORIGIN

  Query Match           100.0%;  Score 858;  DB 6;  Length 858;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-191;
  Matches  858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60

Qy         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120

Qy        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180

Qy        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240

Qy        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300

Qy        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360

Qy        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420

Qy        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480

Qy        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540

Qy        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600

Qy        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660

Qy        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720

Qy        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780

Qy        781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GTTTGTGTAGCCAGTCTACCACCACCACAGCACCCCAGACAGCTAGGCTTAGCCCCTCTG 840

Qy        841 CCCTCCCTTCATTGTACT 858
              ||||||||||||||||||
Db        841 CCCTCCCTTCATTGTACT 858


RESULT 2
AR166506
```

```
LOCUS       AR166506                858 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 2 from patent US 6281190.
ACCESSION   AR166506
VERSION     AR166506.1  GI:16241861
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 858)
  AUTHORS   Hillman,J.L. and Goli,S.K.
  TITLE     Human LEA-motif developmental protein
  JOURNAL   Patent: US 6281190-A 2 28-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .858
                     /organism="unknown"
BASE COUNT      191 a    258 c    245 g    164 t
ORIGIN

  Query Match           100.0%;  Score 858;  DB 6;  Length 858;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-191;
  Matches  858;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCGGTGGTGACTGAGCTACGAGCCTGGCGGCGGGTGTGCGCCGAGCCCCGGCCCGGCCCG 60

Qy         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTCGCGTGCCTCCCAGGCTCCGCACCCCTGATGCTGCGCGGGTGCTGAGCCCGCTTC 120

Qy        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCCGGGACGATGGTGAAGTATTTCCTGGGCCAGAGCGTGCTCCGGAGTTCCTGGGACCA 180

Qy        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGTGTTCGCCGCCTTCTGGCAGCGGTACCCGAATCCCTATAGCAAACATGTCTTGACGGA 240

Qy        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGACATAGTACACCGGGAGGTGACCCCTGACCAGAAACTGCTGTCCCGGCGACTCCTGAC 300

Qy        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAAGACCAACAGGATGCCACGCTGGGCCGAGCGACTATTTCCTGCCAATGTTGCTCACTC 360

Qy        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGTGTACGTCCTGGAGGACTCTATTGTGGACCCACAGAATCAGACCATGACTACCTTCAC 420

Qy        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGAACATCAACCACGCCCGGCTGATGGTGGTGGAGGAACGATGTGTTTACTGTGTGAA 480

Qy        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTCTGACAACAGTGGCTGGACTGAAATCCGCCGGGAAGCCTGGGTCTCCTCTAGCTTATT 540

Qy        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGTGTCTCCAGAGCTGTCCAGGAATTTGGTCTTGCCCGGTTCAAAAGCAACGTGACCAA 600

Qy        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACTATGAAGGGTTTTGAATATATCTTGGCTAAGCTGCAAGGCGAGGCCCCTTCCAAAAC 660

Qy        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTTGTTGAGACAGCCAAGGAAGCCAAGGAGAAGGCAAAGGAGACGGCACTGGCAGCTAC 720

Qy        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGAGAAGGCCAAGGACCTCGCCAGCAAGGCGGCCACCAAGAAGCAGCAGCAGCAGCAACA 780
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 9, 2003, 10:33:38 ; Search time 3494 Seconds
(without alignments)
10045.924 Million cell updates/sec

Title: US-09-892-316-2
Perfect score: 858
Sequence: 1 GCGGTGGTGACTGAGCTACG.....TGCCCTCCCTTCATTGTACT 858

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	858	100.0	858	6	AR028528	AR028528 Sequence
	2	858	100.0	858	6	AR166506	AR166506 Sequence
	3	858	100.0	958	9	BC008866	BC008866 Homo sapi
	4	858	100.0	990	9	BC000007	BC000007 Homo sapi
	5	858	100.0	1006	9	AF111112	AF111112 Homo sapi
	6	858	100.0	1217	9	AF151864	AF151864 Homo sapi
	7	858	100.0	1227	9	BC008307	BC008307 Homo sapi
	8	858	100.0	1243	9	BC007268	BC007268 Homo sapi
	9	858	100.0	1287	9	BC013748	BC013748 Homo sapi
	10	853.6	99.5	1549	6	AR134592	AR134592 Sequence
	11	853.6	99.5	1549	6	AR224016	AR224016 Sequence
	12	844	98.4	956	9	BC013733	BC013733 Homo sapi
	13	841.4	98.1	943	9	AF153607	AF153607 Homo sapi
c	14	811.6	94.6	218836	9	AC016734	AC016734 Homo sapi
	15	810	94.4	193041	9	AC092017	AC092017 Homo sapi
	16	780.4	91.0	1191	9	AF112203	AF112203 Homo sapi
	17	777.4	90.6	40639	9	AL359535	AL359535 Human DNA
c	18	769.6	89.7	181482	9	AC143357	AC143357 Pan trogl
	19	766.4	89.3	1008	9	AF201925	AF201925 Homo sapi
c	20	738.4	86.1	185062	9	AC021231	AC021231 Homo sapi
	21	631.6	73.6	183604	9	AC022022	AC022022 Homo sapi
	22	613.2	71.5	1103	10	BC025859	BC025859 Mus muscu
	23	590	68.8	177054	2	AC126897	AC126897 Rattus no
	24	590	68.8	225563	2	AC134653	AC134653 Rattus no
c	25	569.4	66.4	213940	2	AC110508	AC110508 Mus muscu
c	26	567.6	66.2	200627	2	AC137405	AC137405 Rattus no
	27	567.6	66.2	208649	2	AC110444	AC110444 Rattus no
c	28	567.4	66.1	106019	10	AL627086	AL627086 Mouse DNA
c	29	567.4	66.1	177133	10	AL627074	AL627074 Mouse DNA
c	30	564	65.7	207629	10	AL645637	AL645637 Mouse DNA
	31	535	62.4	593	6	BD108371	BD108371 EST and e
c	32	532.8	62.1	249301	2	AC109524	AC109524 Rattus no
c	33	532.8	62.1	260579	2	AC130181	AC130181 Rattus no
	34	532.8	62.1	288827	2	AC098254	AC098254 Rattus no
	35	513.4	59.8	899	10	BC024813	BC024813 Mus muscu
	36	471.6	55.0	190806	2	AC144798	AC144798 Mus muscu
c	37	464.4	54.1	208265	10	AL606907	AL606907 Mouse DNA
c	38	464.4	54.1	219825	10	AC098886	AC098886 Mus muscu
c	39	430.8	50.2	189667	2	AC123185	AC123185 Rattus no
c	40	427.4	49.8	170952	9	HS501N12	AL022170 Human DNA
	41	423	49.3	872	5	GGU31977	U31977 Gallus gall
	42	421.8	49.2	80457	9	AL157714	AL157714 Human DNA
	43	421.8	49.2	179357	2	AC009625	AC009625 Homo sapi
c	44	414	48.3	108296	2	AC112815_3	Continuation (4 of
	45	414	48.3	181070	2	AC142187	AC142187 Rattus no

ALIGNMENTS

RESULT 1
AR028528
LOCUS AR028528 858 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5858712.
ACCESSION AR028528
VERSION AR028528.1 GI:5940501
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 858)
AUTHORS Hillman,J.L. and Goli,S.K.
TITLE CDNA encoding a LEA-motif developmental protein homologous to avian px19
JOURNAL Patent: US 5858712-A 2 12-JAN-1999;